ABSTRACT OF THE DISCLOSURE

A method and apparatus for data entry by voice under adverse conditions is disclosed. More specifically it provides a way for efficient and robust form filling by voice. A form can typically contain one or several fields that must be filled in. The user communicates to a speech recognition system and word spotting is performed upon the utterance. The spotted words of an utterance form a phrase that can contain field-specific values and/or commands. Recognized values are echoed back to the speaker via a text-to-speech system. Unreliable or unsafe inputs for which the confidence measure is found to be low (e.g. ill-pronounced speech or noises) are rejected by the spotter. Speaker adaptation is furthermore performed transparently to improve speech recognition accuracy. Other input modalities can be additionally supported (e.g. keyboard and touch-screen). The system maintains a dialogue history to enable editing and correction operations on all active fields.